

Practical, Easy-to-Use, Free GIS and Remote Sensing Tools for Resource Management

Andrew Lister
Research Forester
11 Campus Blvd, Ste. 200
Newtown Square, PA 19073
alister@fs.fed.us

Abstract: Geographic Information Systems (GIS) and remote sensing methods have become ubiquitous in natural resource management applications. It is expected that consultants and other practitioners have knowledge of and access to a GIS and remotely sensed data such as digital photographs or satellite imagery. This presentation covers some free, easy-to-use GIS tools that novices can use to learn GIS concepts and to incorporate GIS into their resource management work. Examples of GIS file converters, analytical GIS software, and web-based imagery viewers will be presented in nontechnical language and with an emphasis on their utility for resource measurement and management.

Keywords: Free GIS, free remote sensing software, free GIS data, geospatial analysis, hydrological analysis.

Introduction

Resource managers and environmental consultants often are interested in geographic analyses that will allow them to explore and depict in map format the landscape features of their region of interest. Many also are interested in drawing points, lines or polygons on a map of their area, measuring areas and distances related to these, and characterizing other GIS (Geographic Information System) layers (e.g., elevation, slope or aspect) around them. Finally, users usually want to distribute results of analyses or other geographic information in an appealing, easy-to-interpret way. To perform these activities with little or no cost, one requires 1) a free source and an understanding of appropriate GIS data, 2) free GIS software and conversion tools, and 3) an understanding of basic GIS analysis principles. The goal of this paper is to provide a simple methodology for obtaining and using free GIS information and software, some examples of GIS analyses, and suggestions for presenting the resultant information in an appealing way using free data visualization methodologies.

In this paper, I present an example of a resource analysis exercise using free GIS data and free software for analysis and visualization. I assume that the reader has basic GIS knowledge and a working knowledge of Windows-based computers, including file management, use of the Internet and use of a command prompt for command line programs. Furthermore, I will give references embedded in the text in the form of hyperlinks where the reader can obtain more detailed information to match his or her specific skill level, resource question and GIS data requirements. The references I provide will be entirely Internet-based, allowing resource analysts to obtain the information easily and with limited or no access to a library.

Understanding and Obtaining free GIS data

Free GIS data abound on the Internet in both raster and vector formats. Raster data can be thought of as being like cells in a spreadsheet – each cell has a numerical value for the attribute being mapped, a “pixel size”, or square area on the ground that the cell represents, and is georeferenced (has a connection to a real location on the ground). Maps of the attribute of interest are produced by coloring the cells of a raster data set in a meaningful way (e.g., light to dark colors representing high to low values). Raster data formats include images (e.g., .jpg, .tif, .img, .sid) and other software-specific raster file types (http://www.gis.com/implementing_gis/data/raster.html).

Vector data, on the other hand, consist of georeferenced sets of points, lines or polygons that have tabular, spreadsheet-like attribute data associated with them (http://www.gis.com/implementing_gis/data/vector.html). Each row in the attribute table of a vector format GIS file is usually associated with one point, line or polygon on the map. For example, an individual forest parcel polygon's row (record) can have values for several column variables, like owner code, total forest volume, land use, etc. Vector data are mostly used for identifying boundaries, important geographic features, and areas within or around which additional analyses will be conducted. For example, one could use a vector GIS file depicting stream segments to summarize a raster GIS file representing slope in order to identify areas

within 100 m of streams with steep slopes.

For most resource questions, four GIS data sets are desirable: a vector format file showing the study area boundary and other important point, line, or polygon features; an orthoimage (a raster-based digital photograph that shows the study area in sufficient detail) for examining the vegetation types and other landscape features; a classified landcover map (either raster or vector), where each element in the map has a landcover type assigned to it; and a raster file depicting elevation, which is called a Digital Elevation Model (DEM). With these four files, many useful analyses can be conducted. In the example provided, I will explain how to obtain these data sets, import them into a free GIS, and conduct a simple hydrological analysis – labeling ownership parcels based on their likelihood of contributing to stream sedimentation.

To obtain the vector boundary file and other important vector layers, the user has several options. Many files, such as watersheds, ownership parcels, and political boundaries, can be obtained from a local, state, or federal GIS repository, as described on the National States Geographic Information Council website (<http://www.nsgic.org/states/index.cfm>). If they don't exist, it is possible to turn pairs of x,y coordinates that identify points, lines or polygons into a GIS file. This coordinate information can be manually collected in a spreadsheet using a GPS (Global Positioning System), by manually obtaining it from a paper map, positioning the mouse cursor in free software like Google Earth¹ (<http://earth.google.com>) and adding placemarks that show latitude and longitude values, or from one of the many free mapping sites available on the Internet, e.g., <http://www.infosports.com/m/map.htm> or <http://mapper.acme.com/>.

Once the x,y (generally, longitude and latitude) pairs are assembled into a spreadsheet format, a free, command-line program can be used to create what is probably the most common vector GIS format: an ESRI shapefile (<http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>). Shapefiles consist of three to five files with the same prefix, and some combination of the file extensions .shp, .shx, .sbx, .sbn, .dbf and .prj. All “helper” files must be present for the entire shapefile set to function properly. To create a shapefile from a spreadsheet of coordinate data, one can use the free “Gen2Shp.exe” program¹, which stands for “ArcINFO generate file to shapefile converter”. This program can be downloaded from <http://www.usf.uos.de/~jwagner/Gen2Shp/Gen2Shp-0.2.3.zip>. Once the program is downloaded and unzipped into a directory on your computer (e.g., c:\temp), a command prompt can be opened by choosing “start” then “run” in Windows 2000 or XP, and typing:

```
cmd
```

to invoke the command prompt. Assuming you downloaded the Gen2Shp program to c:\temp, you would then type at the command line in the window that pops up:

```
cd c:\temp
```

and then type the command:

```
Gen2Shp
```

to see the proper syntax of the command required to run the program.

Gen2Shp assumes that your input file is in what's known as an ArcINFO generate file format, a text format that follows the examples shown on <http://intevation.de/~jan/Gen2Shp/> or elsewhere on the Internet. For example, to create a point shapefile called out.shp (and its helper files out.shx and out.dbf), the contents of a text file named in.txt with 3 point coordinates (labeled 1, 2 and 3) would be

```
1,-1.405877,53.721508
2,-2.045479,53.731064
3,-1.984840,53.731972
end
```

and the command to type at the command prompt to create the file would be:

```
Gen2Shp out points < in.txt
```

Gen2shp works similarly with line and polygon format input files, as described on the abovementioned website.

To obtain DEMs, orthoimagery, and other useful data sets, a good source is the United States Department of Agriculture (USDA) Data Gateway. By following instructions on <http://datagateway.nrcs.usda.gov/>, one can either download, or obtain on DVD for a nominal fee, highly resolute, georeferenced, relatively recent imagery, DEM and landcover data covering your area of interest. For my example analysis, I used the 30 m DEM (figure 1) and the 2005 National Agricultural Imagery Program (NAIP) Mosaic orthoimage (figure 2). I also used a subset of the "percent impervious surface" portion of the National Landcover Classification Dataset (NLCD) from 2000 (figure 3), which can be obtained in a similar manner from <http://gisdata.usgs.net/website/MRLC/viewer.php>. Finally, I used existing vector GIS data (tax parcel boundaries) obtained from East Bradford Township in West Chester, PA (www.eastbradford.org) (shown in figures 1, 2 and 3). I recommend carefully noting the latitude and longitude coordinates of the upper left and lower right corners of the analysis area of interest and using them to extract the data sets from the USGS and NLCD data viewers. I also recommend requesting the .tif (geotiff) output format for rasters.

Once these data are downloaded and extracted (the files can be quite large), one needs to determine what is known loosely as the "projection" of the data (http://www.colorado.edu/geography/gcraft/notes/coordsys/coordsys_f.html). Projections of the Earth can be thought of as mathematical transformations of sets of geographic coordinates such that points found on the rounded Earth's surface can be "projected" onto a flat surface, such as a map or a computer monitor. There are several more technical aspects of projections that advanced GIS users need consider, but the novice need only be concerned that the projection method is known, is recognized by the software, and is the same for all data sets being analyzed. Projection information is most often included in metadata files distributed with GIS data.

Free Software and Conversion Tools

Once you have obtained your free geospatial data and know the projection information, you can analyze the data with a free GIS. Several options exist, but one that is well documented and has a lot of useful functionality is ILWIS¹

(http://52north.org/index.php?option=com_projects&task=showProject&id=30&Itemid=127 , and <http://www.itc.nl/ilwis/>). One appealing feature of ILWIS is its ease of installation – all files required are contained in a single directory so the installation of the software simply requires extracting the contents of a zip file to a directory of your choice. In addition, ILWIS performs many DEM-based analyses, which are often useful for land management applications.

Most GIS data sets can be imported to ILWIS format files. The formats for the raster files from the NLCD and USDA archives are .sid and .tif. In order to import these types to ILWIS format files, one could use the native import capabilities of ILWIS, which are based partly on the free Geographic Data Abstraction Library (GDAL) tools for manipulation of GIS data sets. However, the free Frank Warmerdam Toolkit (FW Tools)¹, available at <http://fwtools.maptools.org/> , uses GDAL and allows for more control over many common GIS conversions. After installing the tools, they are accessible via the start menu – the “FW Tools Shell” is a command prompt that opens to the directory in which the tools are stored. For example, to convert a .sid file called 101874.sid stored in c:\temp to an ILWIS format file called 101875IL, one would use the command in the FW Tools Shell:

```
gdal_translate.exe -of ILWIS -projwin xmin ymax xmax ymin c:\temp\101874.sid
c:\temp\101874IL
```

where -projwin xmin ymax xmax ymin are the x and y coordinates for the upper left and lower right corners of the study area you want to define. ILWIS format raster files are approximately 20 times as large as .sid files, so subsetting the files to include just the area of interest is recommended. To convert a .tif file called elev.tif stored in the same directory to an ILWIS file called elev, use the command:

```
gdal_translate.exe -of ILWIS -projwin xmin ymax xmax ymin c:\temp\elev.tif
c:\temp\elev
```

Conveniently, ILWIS recognized the projections of the files that were imported using FW Tools from the USDA archive (in this example, both are UTM zone 18, NAD83). The NLCD data set, however, was in Albers NAD83, and needed to be projected to match the UTM data. Using the ILWIS “resample” operation, which is used to change the projection of rasters, I converted the NLCD Albers projection to the same UTM projection as the DEM and the NAIP imagery by assigning it the exact same projection possessed by the DEM, as described in the ILWIS help.

To import the shapefile of tax parcels, I used the ILWIS “Import Map” function. ILWIS did not seem to recognize projection information associated with the shapefile, so, using the methods described in the ILWIS help, I created a coordinate system in ILWIS, defined it as latitude/longitude, decimal degrees, NAD83 (which was the coordinate system of the original shapefile), and assigned it to the newly imported ILWIS vector file. Once I correctly imported the file and defined its projection, I “transformed” it, or changed its projection to match that of

the DEM, by using the “Vector Operations/Transform Polygons” function and choosing the DEM's projection definition as the output. Transforming and resampling can be very time intensive, so I recommend avoiding them, especially resampling, whenever possible, and choosing the coordinate system possessed by the majority of the data sets as the basis for all analyses.

Understanding basic GIS principles

I will present an example analysis using the parcel boundaries, the DEM, and the NLCD landcover map to rank parcels based on their potential for contributing sediment to streams. I will first preprocess my DEM to prepare it for a procedure in which I create a drainage network map. I will then process this drainage network map to create a map of overland flow distance from each pixel to the nearest drainage. With this information, a slope map, and the NLCD percent impervious cover maps, I will derive an index whereby each pixel will be assigned a composite value that indicates its potential for contributing sediment to the drainage network. I will then label each ownership parcel with the average index value, thus creating a map of ranked parcels that can be used for management purposes. Throughout the discussion, I will give some brief background on GIS usage principles that are *a propos* to the analysis being conducted.

To conduct a hydrological analysis, the DEM being used needs to be as flawless as possible. Many DEMs have errors or anomalies that can be fixed using ILWIS's “DEM Hydroprocessing/Flow Determination/Fill Sinks” operation. Sinks are pixels that do not “fit in” with their neighbors – anomalous dips in the landscape that could affect GIS analyses that rely on calculating cell-to-cell flow of water. Once this preprocessing operation was complete, I performed the “Flow Direction” (figure 4) and “Flow Accumulation” (figure 5) operations. These operations assign to each pixel, respectively, the direction that water would flow naturally as it passed over that pixel, and the cumulative number of pixels that would naturally drain into a given pixel. While useful in and of themselves, these outputs were part of the preprocessing steps prior to extracting the drainage network.

The output of the “flow accumulation” raster was used as input for the “Drainage Network Extraction” operation, which served to create a binary map in which pixels were labeled as “true” if they are part of a drainage system, or “false” if they are not (figure 6). I had the option of specifying the minimum area that must flow into a pixel for it to be considered a drainage, and I chose an area of 20 pixels, equivalent to about 5 acres. Although this level of detail was desirable because it was relevant to the scale at which ownership parcels might be managed, it did introduce anomalies into the analysis that I was unable to eliminate, particularly in flat areas (figure 7). In future studies, I could choose to do some manual editing in these areas to eliminate this problem, but for the current example, the affected areas were near the periphery of the study area, so I decided to simply make note of the presence of inaccuracies and where they occurred.

To implement the “Overland Flow Length” operation, I assigned stream order to each segment of my drainage network. This was done using the “Drainage Network Ordering” operation. The drainage network order raster was used as input to the “Overland Flow Length” tool, which produced the map shown in figure 8 – each pixel is assigned the distance that water flowing

naturally from it would have to travel before it reached a drainage. To prepare the map for compositing with the other layers comprising my sedimentation risk index, I standardized the map values to range between 1 and 100, with 1 being the largest distance and 100 being the shortest distance (because being a short distance from a drainage makes a pixel more likely to contribute sediment) to a drainage. I performed this operation using the “Map Calculation” tool. If you conceptualize a raster data set as being like a spreadsheet, the Map Calculation Tool is like an equation editor that applies an equation to each cell of a spreadsheet and returns values in a set of cells with the exact same extent as the original spreadsheet. The equation is entered into the interface, and it is applied on a cell-by-cell basis across the entire file. The equation used was:

$$(1 - ('flowlength' / \text{MapMax}('flowlength'))) * 100$$

where MapMax is the maximum value found on the map, and 'flowlength' is the name of the output of the “Overland Flow Length” operation.

Next, I calculated slope percent, as described in the ILWIS help files, to produce a map of the percent slope by calculating elevation differences between adjacent pixels. I standardized this slope map in a similar way, using the equation:

$$(((\text{MapMax}('slope') - 'slope') / \text{MapMax}('slope'))) * 100$$

where 'slope' is the percent slope map. This formed the second component of my sedimentation risk map.

Finally, I added these standardized maps to the NLCD percent impervious map (which already had values ranging from 0 -100) using the “Map Calculation” equation:

$$'flowlength_std' + 'slope_std' + 'imperv'$$

where flowlength_std and slope_std are the standardized flowlength and slope outputs, and imperv is the percent impervious layer, to obtain figure 9, the sedimentation risk index map. Values closer to 300 had relatively steep slopes, were close to a drainage, and had a lot of impervious surface. Values closer to 0 had a large amount of vegetative cover, were relatively flat, and were far from drainages. It would be very easy to weight the layers differently or choose a different combination of factors to create the index, but that is beyond the scope of this paper.

As a last step, I intersected the sedimentation risk raster with the parcel boundaries layer by first rasterizing (converting to a raster format) the parcels layer using the “Rasterize/Polygon to Raster” operation, and then using the “Cross” operation, which creates a table showing every unique combination of parcel identification (ID) value and sediment index value that occurs in any one location on the map. By adding a new column to this table using the table editor, and then assigning the average value of the sediment index values using the table editor's “Columns/Aggregate” operation and choosing to group by parcel ID value, I computed the average sedimentation index value within each parcel. I then assigned this table to be the attribute table of the “Cross” operation's output raster, and label the final map with average

sedimentation risk by parcel (figure 10).

Displaying Results in Innovative Ways

The Google Corporation provides the Google Earth software (<http://earth.google.com>) free of charge. In addition to providing a 3-D view of relatively current satellite and orthoimagery, it allows for the display of user-supplied, georeferenced vector data sets in the .kml file format. There are several ways to convert ILWIS point, segment, or polygon layers to .kml format. For example, the FW Tools suite has a program called ogr2ogr.exe, which converts shapefiles to several formats, including .kml, with the following command:

```
ogr2ogr -f "KML" -t_srs EPSG:4326 outfile.kml infile.shp
```

where “outfile” is the desired kml filename, and “infile” is the shapefile that was exported from ILWIS using the “Export” operation, which is described in the ILWIS help. This method is simple and fast, but the resulting kml file has only limited color options, and thus little facility for depicting variations in attribute values. This command would be most appropriate for exporting files like a drainage network vector data set to view in Google Earth (figure 11). There are several other shapefile to kml converters that are very inexpensive (e.g., tiles2kml¹ (<http://tiles2kml.com>), and are worth investigating. For a similar method of displaying output shapefiles using Google technology, the GMapCreator software¹ (<http://www.casa.ucl.ac.uk/software/googlemapcreator.asp>) allows for publishing shapefiles to the Internet, changing their colors to highlight certain attribute values, and overlaying them on satellite and other imagery.

A second option, which allows for more interactive exploration of analysis results, is to use one of the free GIS data viewers available on the Internet. Mapwindow GIS¹ (<http://www.mapwindow.com/>) is free and it allows not only for display of GIS data, but for many complex analyses of raster files, rivaling those offered by ILWIS. ESRI provides the popular ArcExplorer software¹ (<http://www.esri.com/software/arcexplorer/explorer.html>), which is very easy to use and allows for connections to web map services (WMS). Project files that save the colors, legends and other features of map displays can be saved and distributed along with the software so data consumers can see analytical results in the exact way that you chose to format them (assuming map users are willing to install ArcExplorer).

Summary

Resource managers have at their disposal a large amount of GIS data, free software and conversion tools, and useful Internet sites and other documentation that can provide the tools needed to conduct complex analyses and display results in an appealing way. The advantages of using free GIS software and data are many, but there are some disadvantages, including lack of support for some software problems, unknown software providers, and in some cases, poor documentation. However, with software like ILWIS, the simple methods and explanations provided here, and a willingness to learn, natural resource professionals can improve their ability to understand and analyze their region of interest using geospatial technology.

¹ The use of trade, firm or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the U.S. Department of Agriculture or the Forest Service of any product or service to the exclusion of others that may be suitable.

Figures

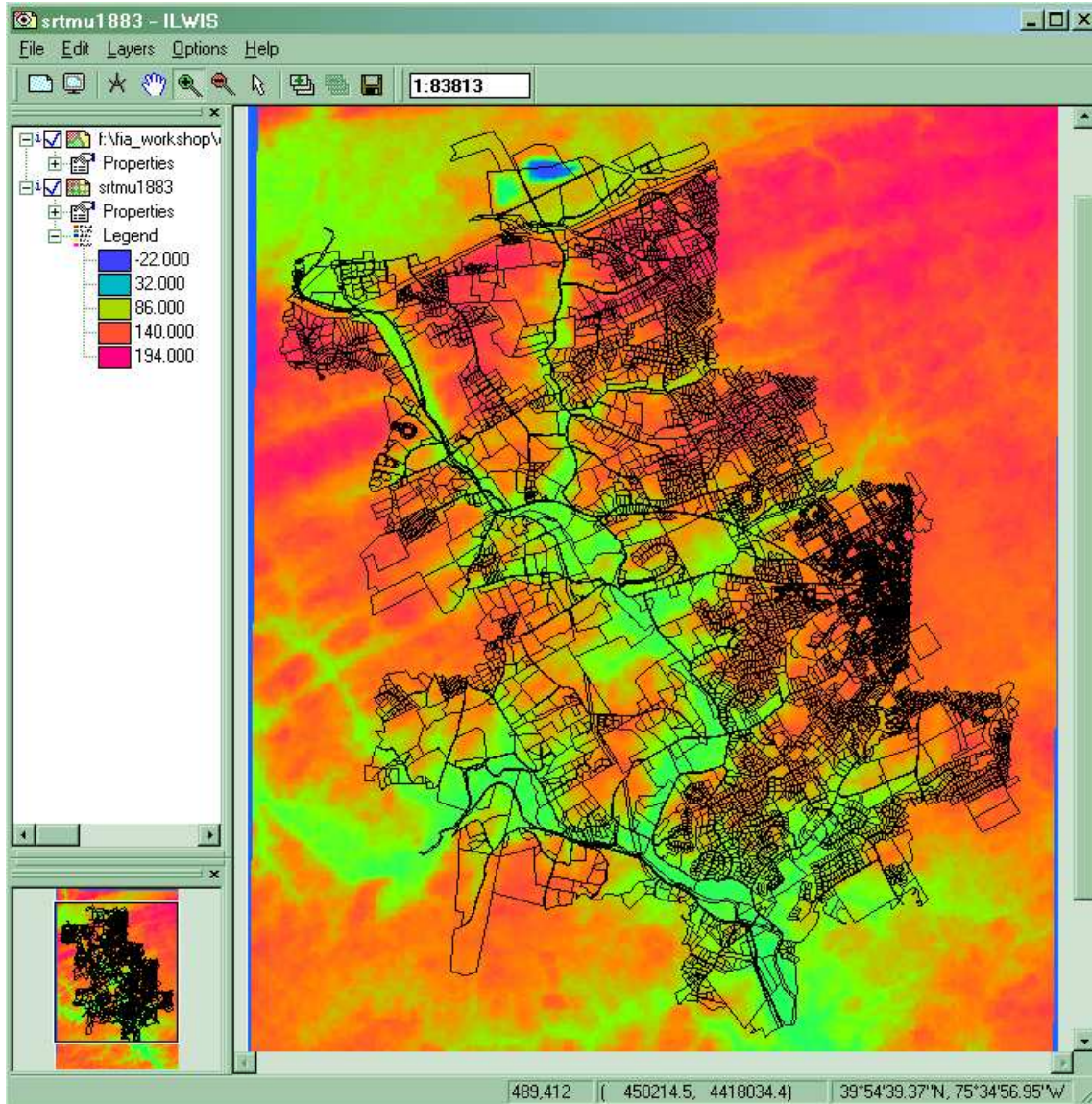


Figure 1. The DEM obtained from the USDA Data Archive website, with ownership parcels for East Bradford Township overlain, displayed in the ILWIS software window.

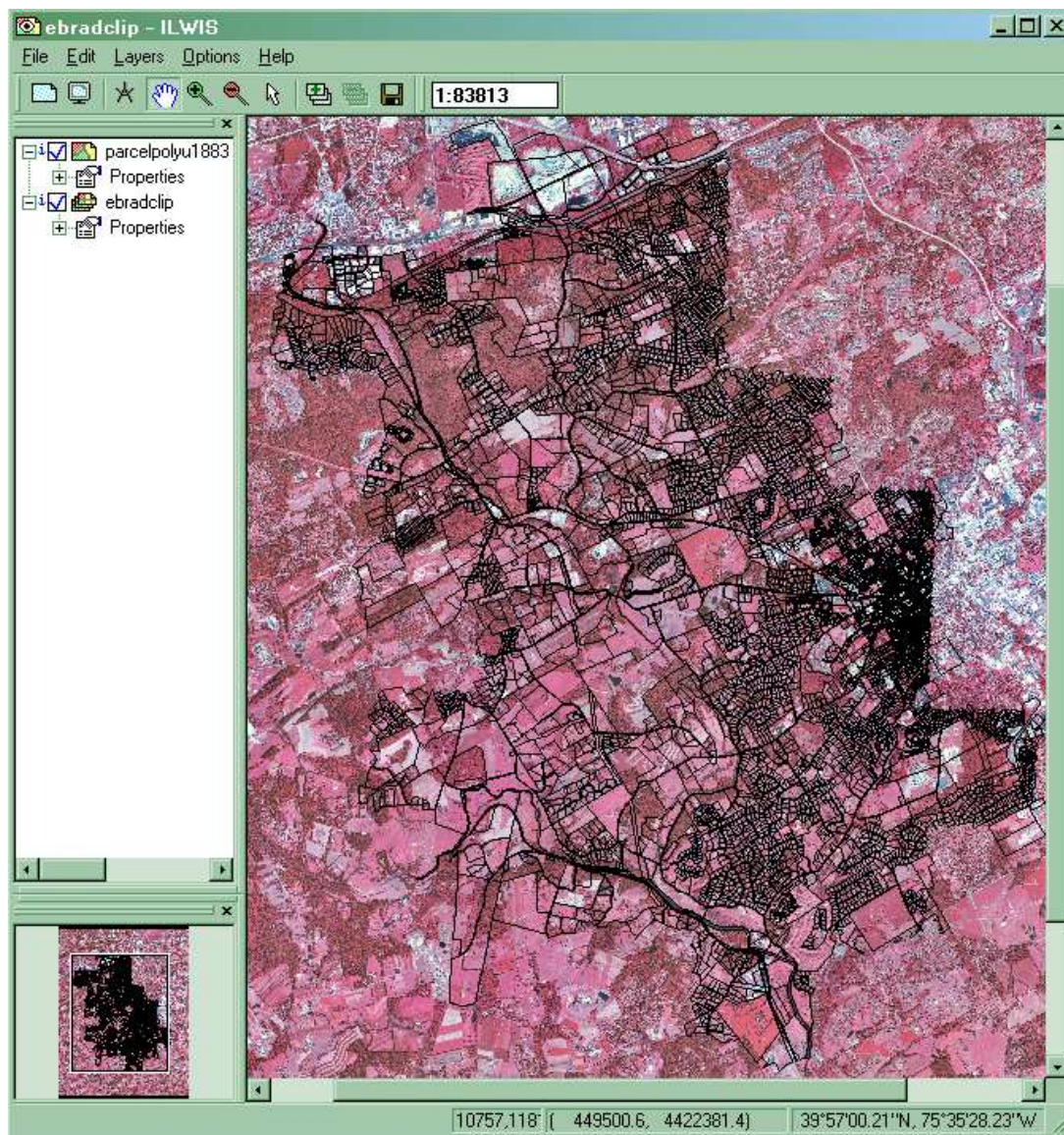


Figure 2. The NAIP orthoimagery obtained from the USDA Data Archive website, with ownership parcels for East Bradford Township overlain, displayed in the ILWIS software window.

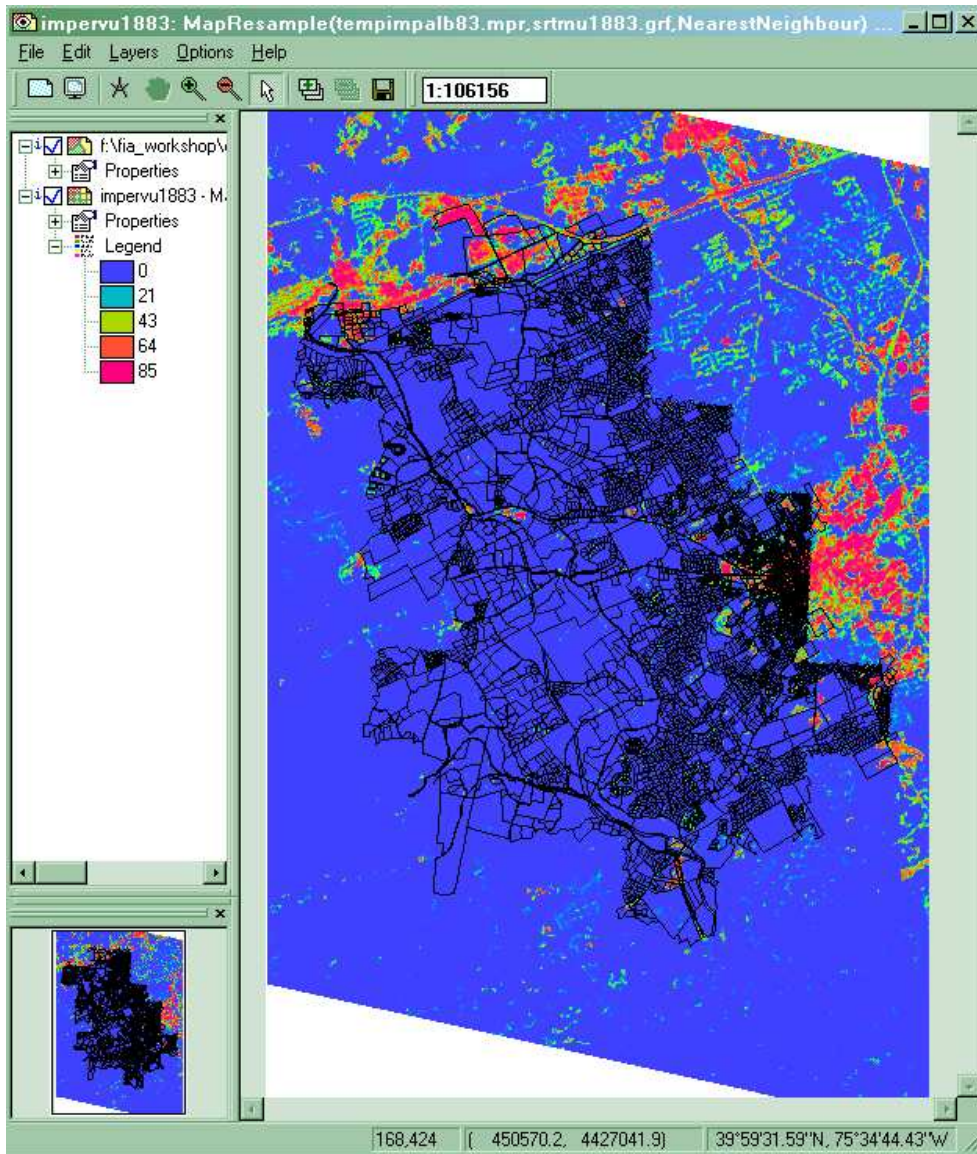


Figure 3. The NLCD percent impervious data set obtained from the NLCD Seamless Data Distribution website, with ownership parcels for East Bradford Township overlain, displayed in the ILWIS software window.

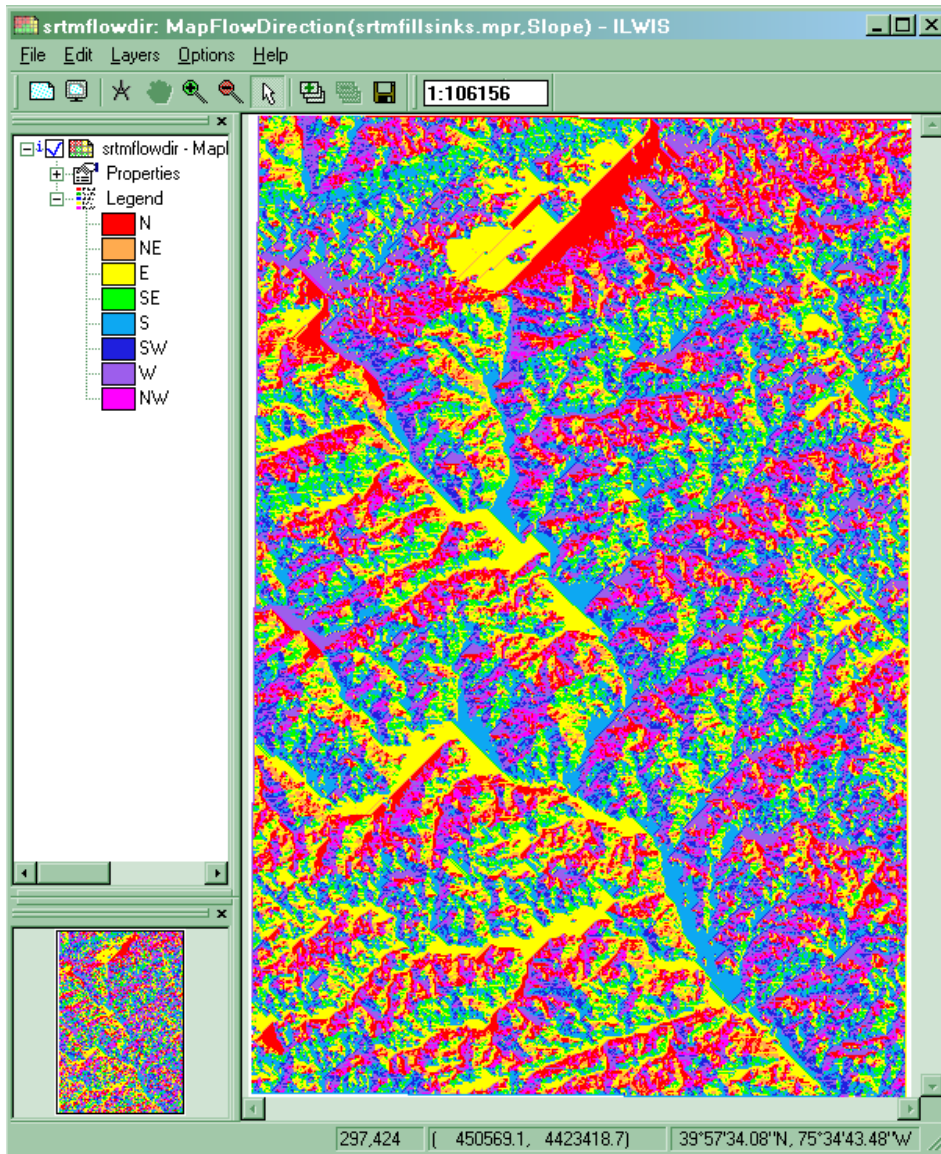


Figure 4. The output of the “Flow Direction” operation. Each pixel is assigned the cardinal direction toward which water would naturally flow on the landscape.

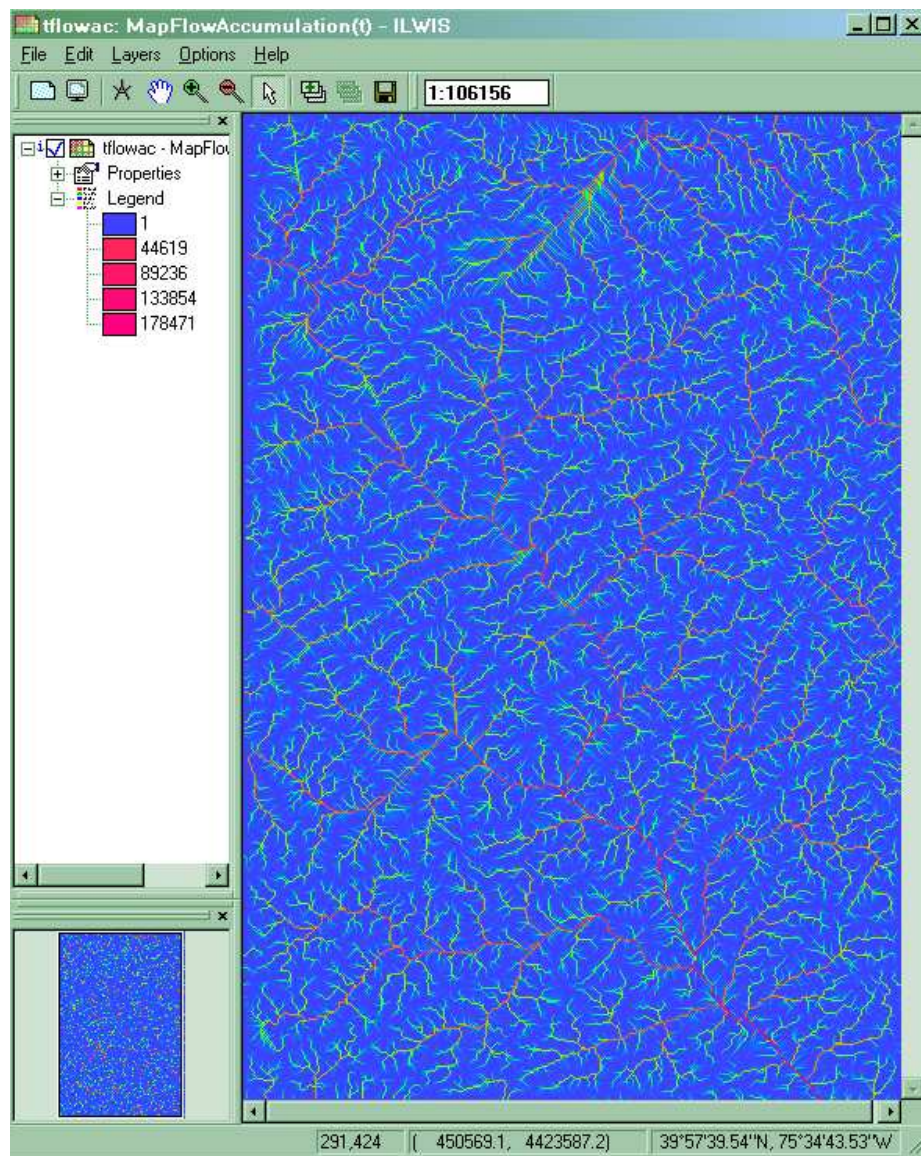


Figure 5. The output of the “Flow Accumulation” operation. Each pixel is assigned the cumulative count of all pixels that would naturally flow into it.

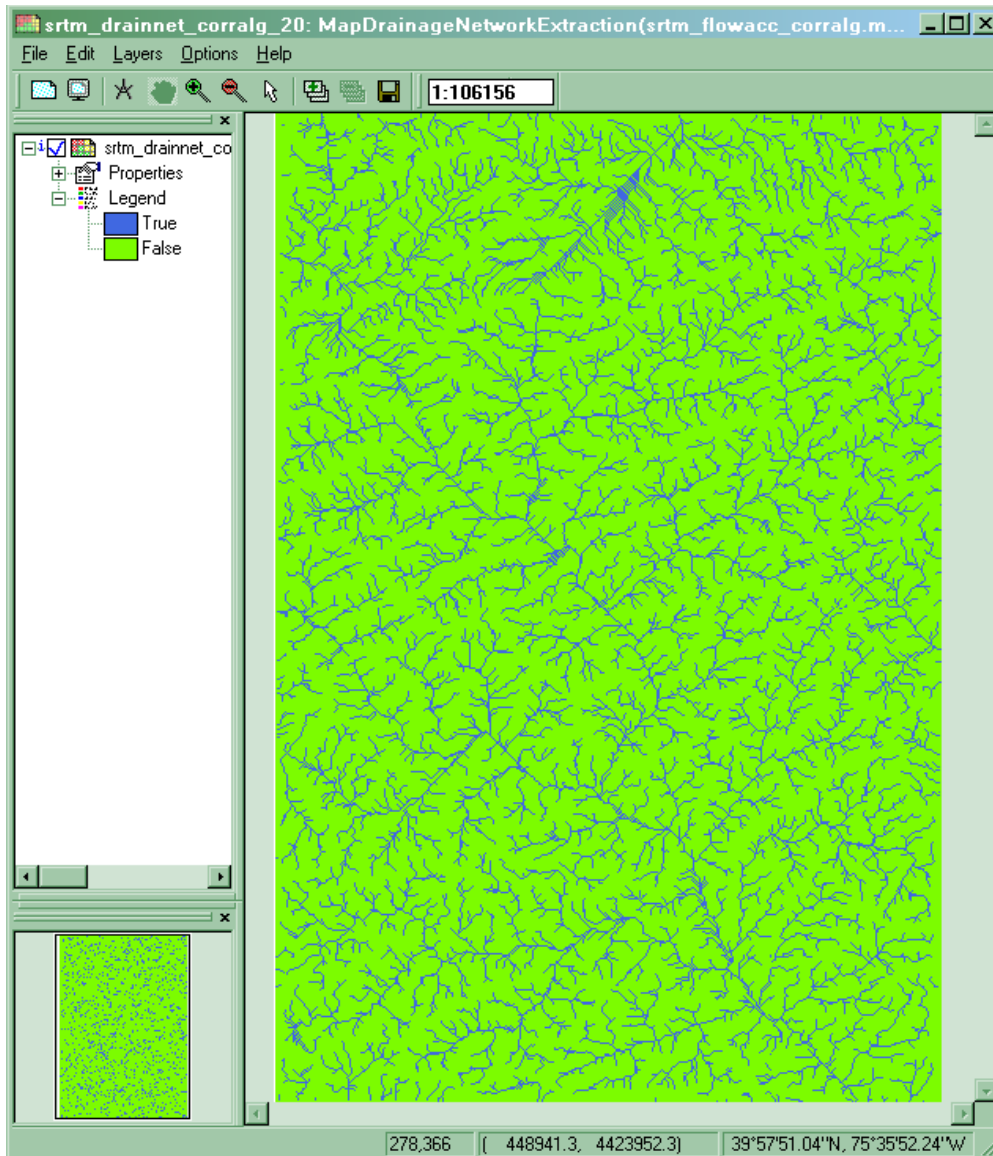


Figure 6. Output of the “Drainage Network Extraction” operation. A pixel can be considered a drainage if at least 20 other pixels drain into it.

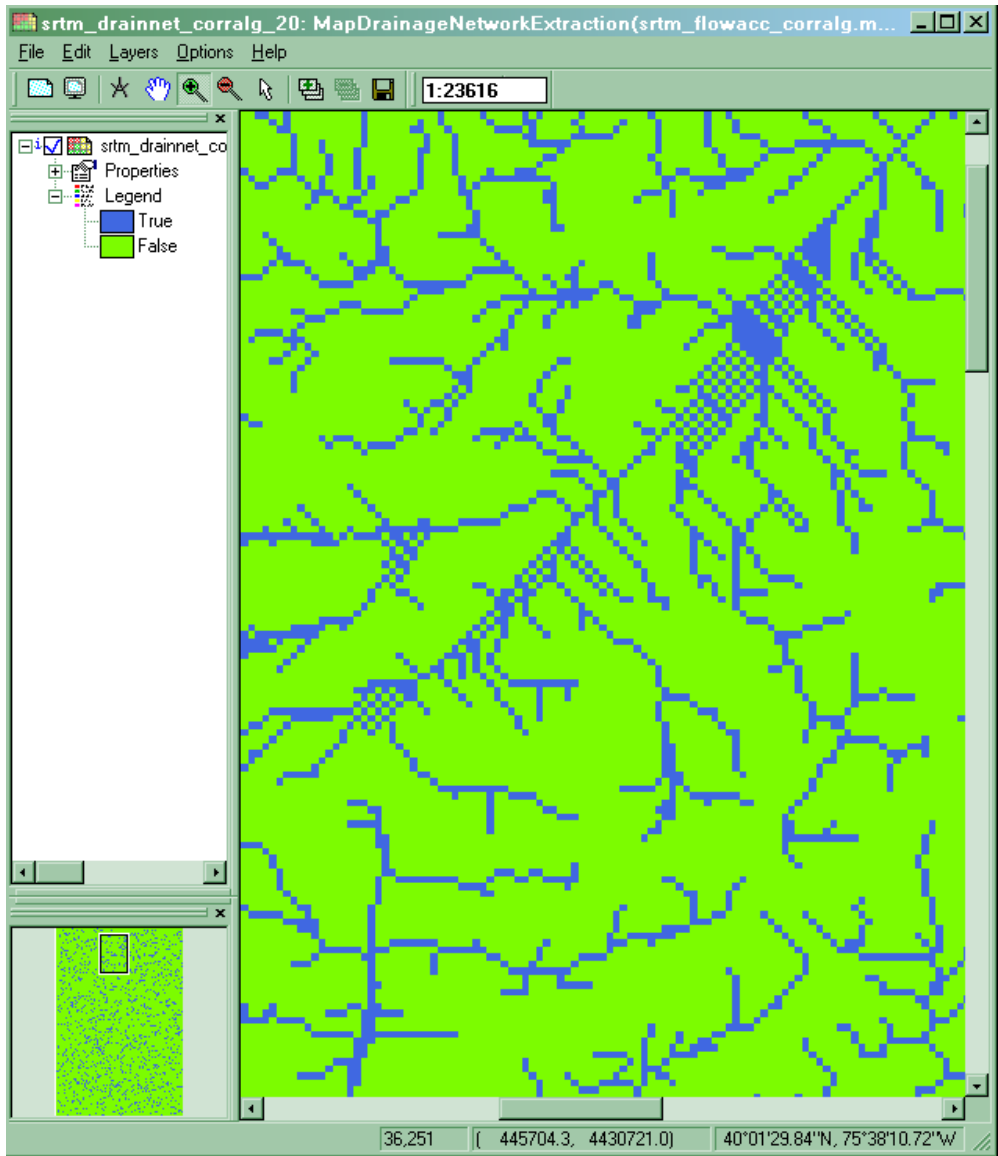


Figure 7. Example of an anomaly in the drainage network caused by large, flat regions in the DEM. Manual editing could be used to fix errors like these.

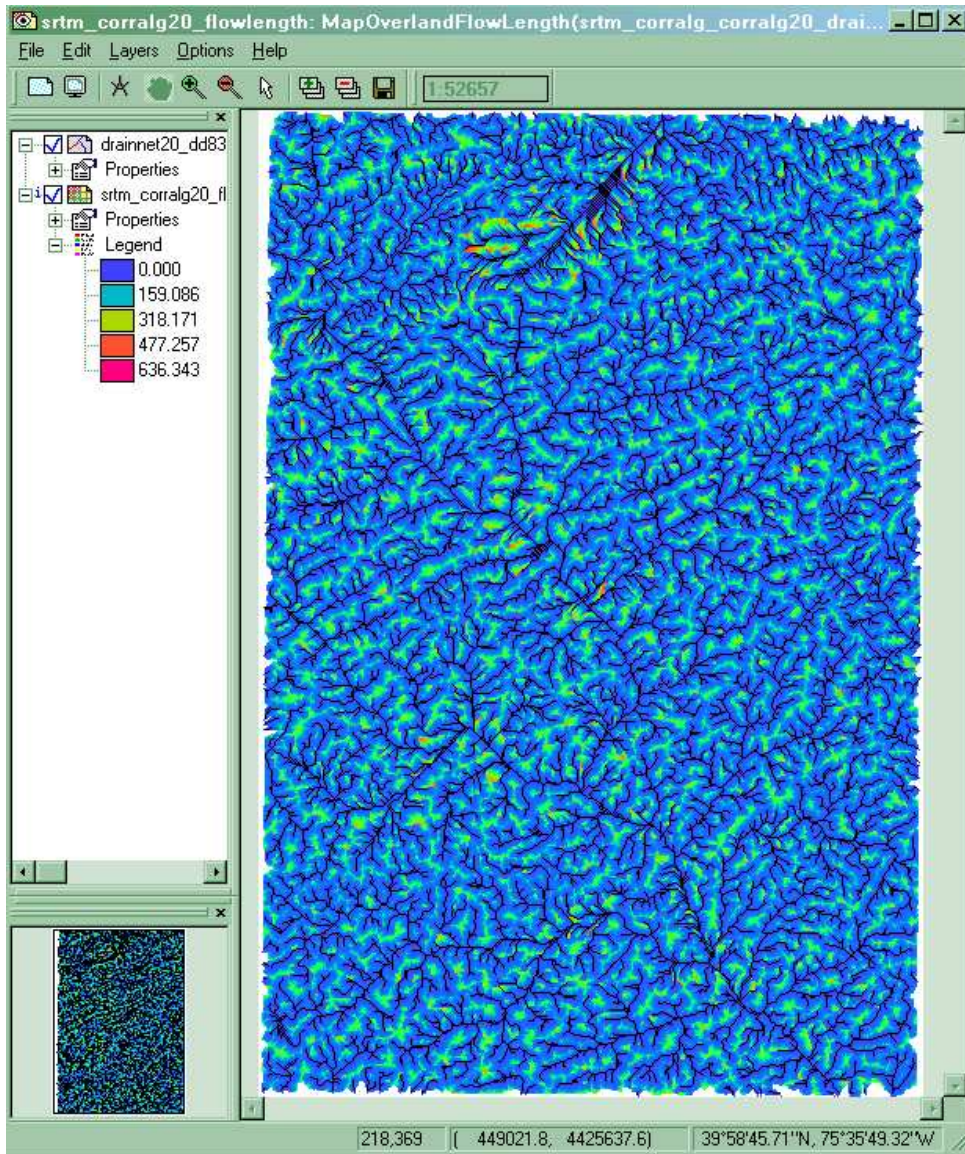


Figure 8. An example of the output of the “Overland Flow Length” operation, which calculates how far water flowing over each pixel must travel before reaching a drainage.

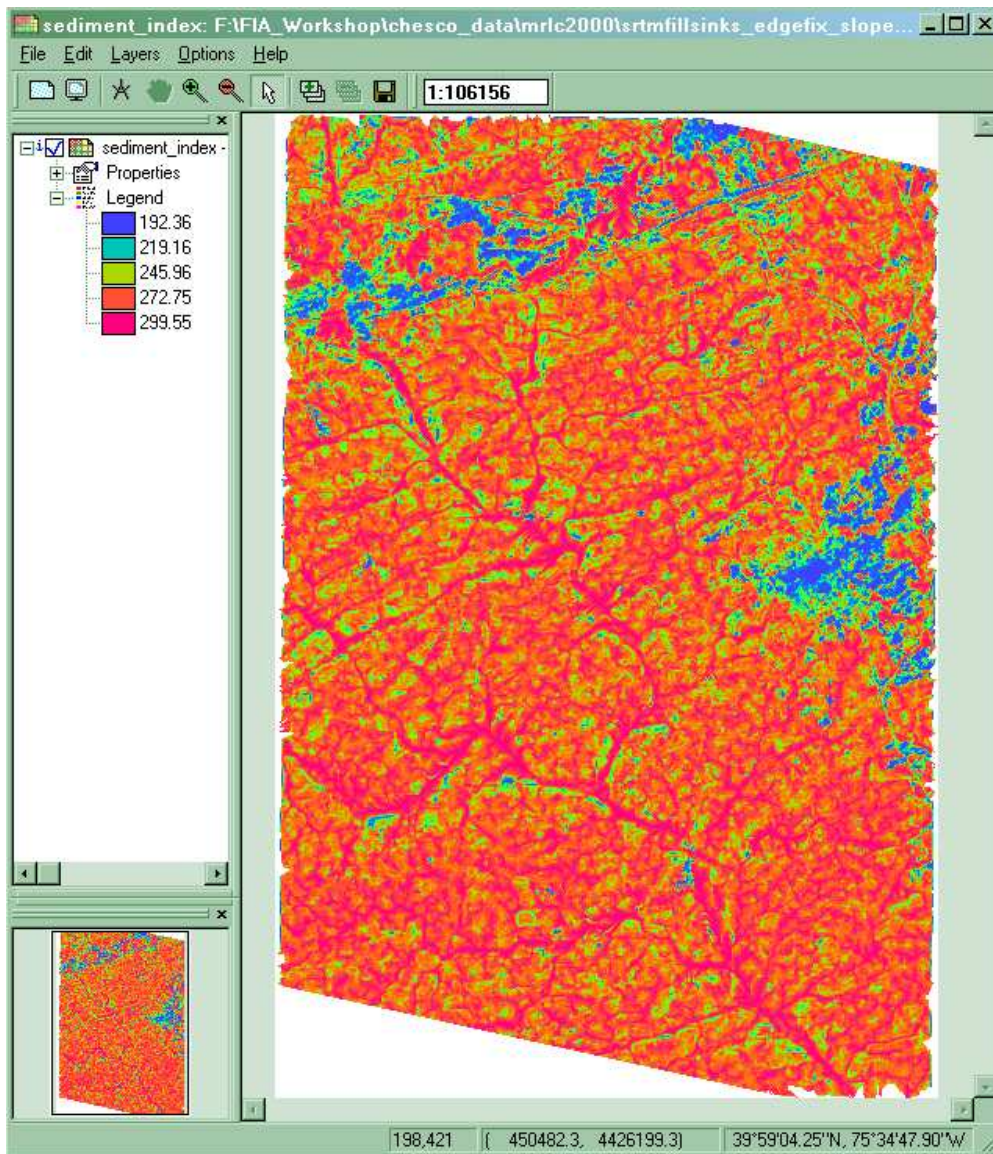


Figure 9. Sedimentation index map. Each pixel contains the sum of standardized slope, overland flow length, and percent impervious layers. High values (near 300) indicate high sedimentation risk for the drainages that include the pixels, and low values represent low sedimentation.

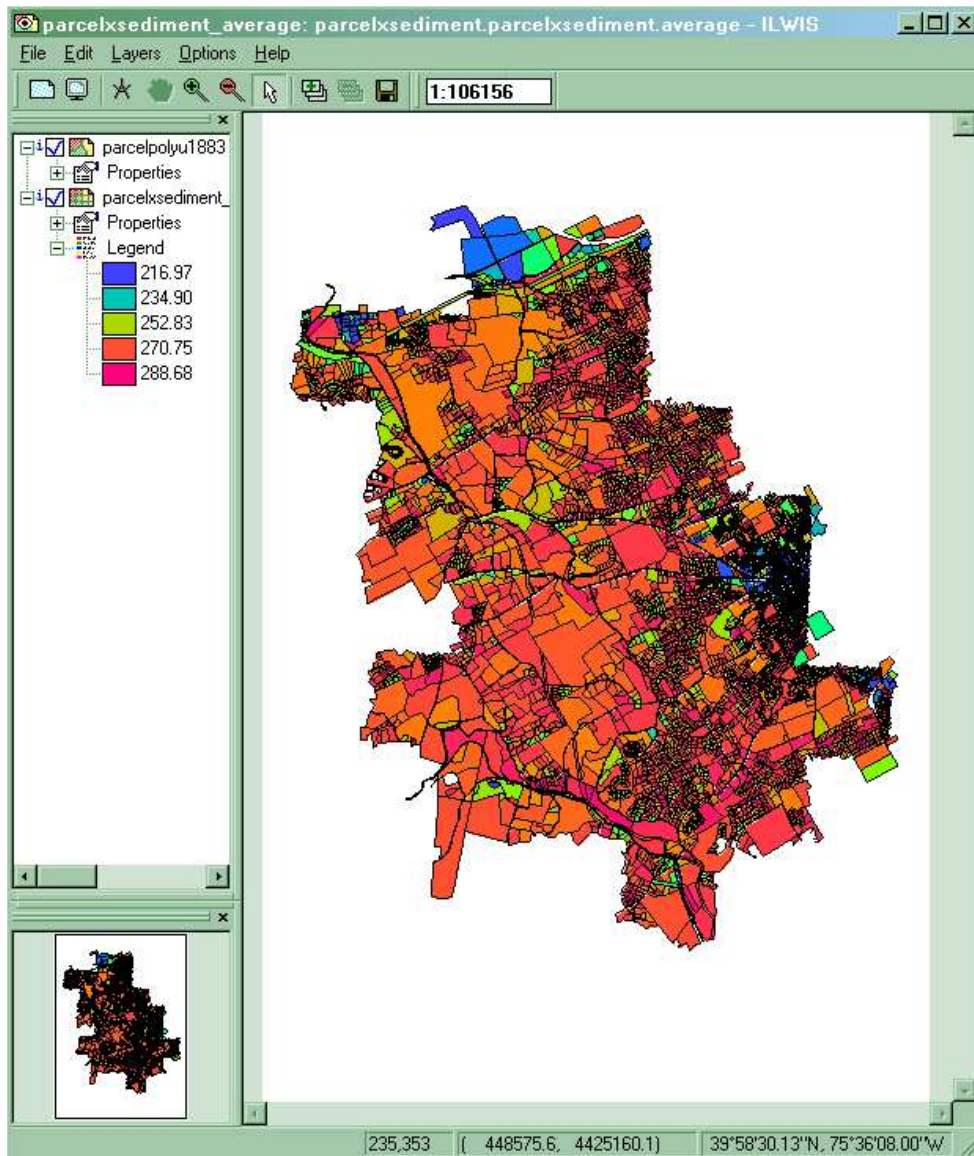


Figure 10. Sedimentation risk index aggregated to the ownership parcel level. Each parcel is assigned the average risk value within its boundaries.

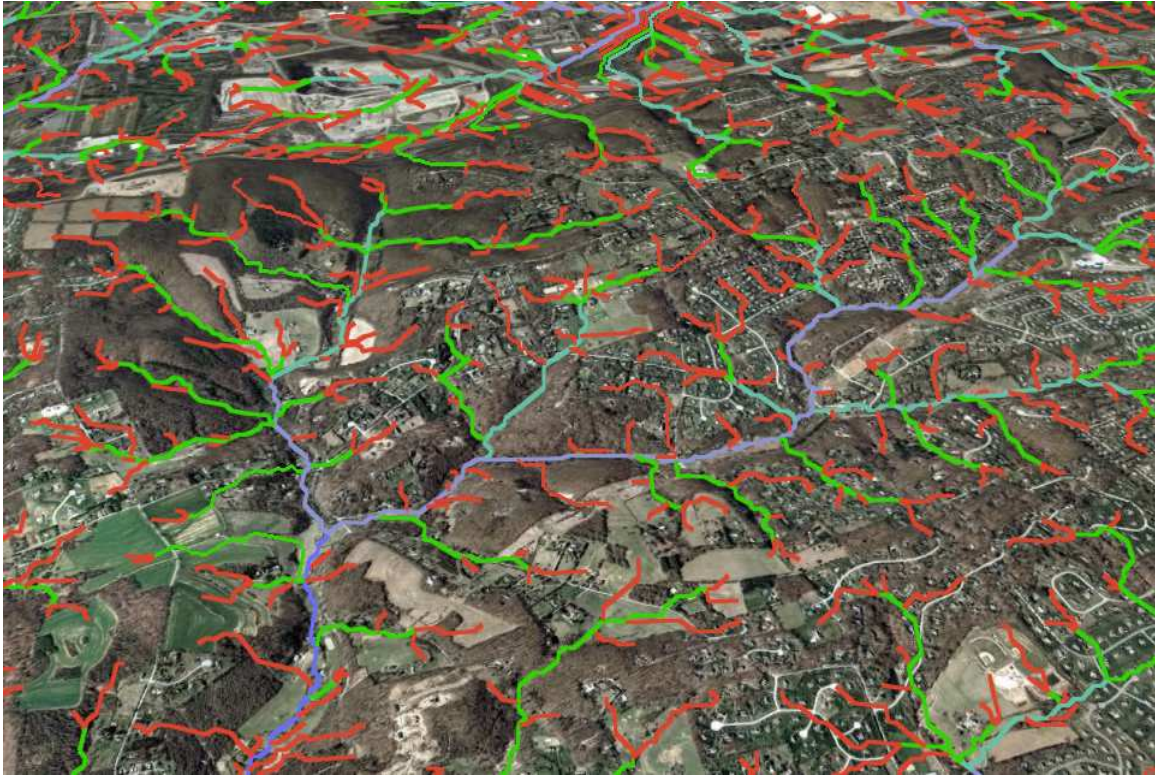


Figure 11. The kml output of the ogr2ogr.exe operation, depicting a drainage network overlaid on an orthoimage of the study area in Google Earth.